Distinct functions of genome-wide chromatin remodeling in the differentiation of T helper Type 1 and 2 cells

Wei-ping Zeng (✉ weipingzengny@gmail.com)

Marshall University

Research Article

Keywords:

Posted Date: December 13th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1164286/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Title: Distinct functions of genome-wide chromatin remodeling in the differentiation of T helper Type 1 and 2 cells

Wei-ping Zeng^{1,2,3}

Departments of Biochemistry and Microbiology¹ and Pediatrics², J. C. Edwards School of Medicine; Center for Cell Development and Differentiation, Department of Biology, College of Sciences³, Marshall University, 1700 3rd Avenue, Huntington, WV 25755

Correspondence and lead contact: Wei-ping Zeng, E-mail: weipingzengny@gmail.com

Summary

T helper type 1 and 2 (Th1 and Th2) cells play critical roles in infectious, autoimmune and allergic diseases. Here we mapped genome-wide distribution of DNase I hypersensitive (DHS) sites in Th1, Th2 and their precursors naïve CD4 T cells. DHS sites were found unevenly distributed in the genomes with highest densities within 2kb of the transcription start sites (TSS). At the whole genome level, the DHS values, representing chromatin openness, but not the numbers of DHS sites showed strong positive correlation with gene expression. Th1 and Th2 differentiations were accompanied by changes of genome-wide distribution of DHS sites. The differentiated cells assumed more open chromatin structures than their precursors. During Th1 differentiation changes of DHS values could be statistically positively or negatively associated with changes of gene expression depending on the locations of the DHS sites, whereas only positive association was found during Th2 differentiation.

Introduction

Successful clearance of different infections requires tailor-made immune responses to the specific infectious agents. To accomplish this, conventional naïve CD4 T cells differentiate into different effector T helper subsets, each of which produces a distinct set of cytokines that mediate their effector functions. The first described subsets are the T helper type 1 and type 2 (Th1 and Th2) cells (1). The Th1 cells produce their signature cytokine IFN- γ , whereas the Th2 cells produce IL-4, 5 and 13. Th1 responses are required for the clearance of intracellular bacterial infection (2), and Th2 responses are required for the clearance of helminth infections (3). While the protective immunities provided by Th1 and Th2 responses to infections benefit the hosts, Th1 and Th2 responses to self-antigens and allergens cause detrimental immunopathology in organ specific autoimmune and allergic diseases, respectively (4, 5).

Differentiation of naïve CD4 T cells into Th1 or Th2 cells is induced by the combination of TCR stimulation and polarizing signals. IL-12 and IL-4 provide strong polarizing signals for Th1 and Th2 differentiation, respectively (6, 7). Other cytokines such as IL-33 and TSLP, as well as Notch ligands on the cell surface, also provide polarizing signals (8-11). These signals cooperate to induce Th1 and Th2 specific master transcription factors to further drive Th1 and Th2 differentiations, which are mutually exclusive (12, 13). Once differentiated, Th1 and Th2 cells gain the ability to produce their effector cytokines upon re-stimulation by the TCR without the original polarizing signals.

Chromatin remodeling is a widely recognized feature of cell differentiation. Chromatin remodeling changes the distribution of “open” chromatin regions, which are often referred to as DNase I hypersensitive (DHS) sites for their high sensitivity to digestion by DNase I (14). The open chromatin structures at the DHS sites allow transcription factors to access *cis* regulatory sequences. In fact, various *cis* regulatory elements such as promoter, locus control region (LCR), enhancer, boundary element and even silencer are located at the DHS sites (10, 15, 16).

Previous studies of chromatin remodeling during Th1 and Th2 cell differentiation have focused on DHS sites at the genomic loci of the effector cytokine genes. Subset specific DHS sites are detected at the Th2 cytokine and IFN- γ gene loci, and they play critical roles in the specific expression of the effector cytokines (10, 17-19). However, it remains unclear whether chromatin remodeling occurs only at the local levels or there is genome wide chromatin remodeling during the Th subset differentiation. In the present study, we compared genome-wide distribution of DHS sites in naïve CD4 T cells, Th1 and Th2 cells, and analyzed the relations between genome wide chromatin remodeling and global gene expression profiles.

Results

Genome-wide identification of DHS sites

To gain understanding of genome-wide distribution of open chromatin regions, we identified DHS sites in Th1, Th2 and their precursors naïve CD4 T cells using the DNase-Seq protocol developed by Boyle et al (20). Figure 1 shows representatives of the 140bp PCR products of the DHS libraries that contain 20bp of genomic DNA sequences derived from the DNA ends generated by DNase I digestion of the nuclei of naïve CD4 T cells, Th1 and Th2 cells. For each cell type, multiple independent replica libraries were constructed. The library read counts ranged from 21.3 to 60.2 millions. The reads were mapped to the mouse genome, and DHS peaks were called using F-Seq (21). DHS peaks of a representative library at the *CD4*, *IFN- γ* and *IL-4/13* loci are shown for naïve CD4, Th1 and Th2 cells, respectively (Fig S1).

Characteristics of genome-wide distribution of DHS sites

We assumed that the location of a DHS site is a key determinant of whether the DHS site plays a role in regulating its associated gene. We therefore analyzed the distribution of DHS sites at distances relative to their associated genes. We plotted the numbers of DHS sites against the distances to the transcriptional start sites (TSS) or transcriptional termination sites (TTS) of their associated genes. Large numbers of DHS sites were concentrated near the TSS, and to lesser degrees near TTS (Fig S2B, C). To gain better resolutions, we limited the genomic distances to cover only 80% of the DHS sites. As shown in Fig 2B, sharp spikes of DHS peaks were found near the TSS in all cell types. The numbers of peaks decreases as the distance to TSS increases in both the up- and downstream directions. Similarly, the numbers of DHS sites decrease when the distances to the TTS increase (Fig 2C). Based on these observations, we divided the DHS sites into 6 location categories, including -5kb to -2kb from TSS, -2kb to TSS, TSS

to +2kb, +2kb to TTS, TTS to TTS +5kb, and intergenic regions (Fig 3, left). To better show the distribution of the DHS sites in genome wide landscape, we calculated DHS peak densities (peaks/kb) in each location category. Highest peak densities were found in the -2kb to TSS and TSS to +2kb categories. The regions with the second highest peak densities were the regions from -5kb to -2kb upstream of the TSS and within 5kb downstream of TTS. On the other hand, the regions from +2kb to TTS and the intergenic regions showed relatively low peak densities (Fig 3, right).

Relations between DHS and gene expression in each cell type

To quantitatively measure the “openness” of the chromatin structure of a DHS site, we assigned a DHS value to each DHS site as the read counts of 600bp genomic interval centered at the position of the DHS peak normalized to the average read counts of all DHS sites in a cell type (22). On the other hand, gene expression level was determined as RPKM. We divided the genes in each cell type according to their RPKM values (top 25%, 50% and bottom 50% and 25%), and compared the DHS values of the DHS sites associated with the genes. When all DHS sites were analyzed regardless of their locations relative to the genes, DHS values decreased along with RPKM (Fig 4A). When the DHS sites in different location categories were examined, strong positive correlations between DHS values and RPKM were found in DHS sites in the location categories of -2kb to TSS and TSS to +2kb, whereas the correlation was much weaker with DHS sites of other location categories (Fig 4C). In contrast to the DHS values, the total numbers of DHS sites showed only marginal, if any, positive correlation with the RPKM of their associated genes (Fig 5A). Nonetheless, the positive correlation was more notable in the +2kb to TSS and intergenic location categories than other categories (Fig 5C).

Correlation coefficients between DHS and RPKM values

As an independent method of determining the correlation between DHS values and RPKM, we calculated Kendall rank correlation coefficients (tau values). In agreement with Fig 4A, positive correlation coefficient values were obtained for all 3 cell types when all DHS sites were considered (Table 1). Similarly, highest positive correlation coefficients were obtained with DHS sites in the -2kb to TSS and TSS to +2kb location categories (Table 1).

Dramatic changes of DHS site during Th1 and Th2 differentiations

To determine whether there is genome wide chromatin remodeling during the differentiation of Th1 and Th2 cells from naïve CD4 T cells, we compared DHS sites in naïve CD4 T cells with those in Th1 or Th2 cells. We calculated the changes of DHS, i.e., Δ DHS, values when naïve CD4 T cells differentiated into Th1 or Th2 cells (22) for each of the DHS sites identified in the naïve CD4 T cells and Th1 or Th2 cells. As shown in Fig 6, during Th1 or Th2 differentiation from naïve CD4 T cells only small fractions of DHS sites remained unchanged. Large numbers of DHS sites had either increases or decreases in DHS values. The DHS sites with increases in DHS values far outnumbered those with decreases in DHS values, demonstrating that overall the differentiated cells assumed more open chromatin structures than their naïve precursors.

Correlations between the changes of DHS values and gene expression during Th1 and Th2 differentiation

Kendall rank correlation coefficients were calculated between Δ DHS of DHS sites and the changes of expression levels, i.e., Δ RPKM, of their associated genes (Table 2).

When all DHS sites were considered, we detected positive correlations between Δ DHS and Δ RPKM during Th1 and Th2 differentiation. When DHS sites were divided into

different location categories, strongest positive correlations were detected in the categories of -2kb to TSS and TSS to +2kb. For the remaining categories during Th1 differentiation, 1 category (TTS to TTS + 5kb) showed somewhat weaker positive correlation, but more strikingly 2 categories (+2kb to TTS and intergenic) showed statistically significant negative correlations; and 1 category (-5 to -2kb) showed negative correlation that was not statistically significant. The negative correlations suggest that overall DHS sites in these location categories may be more likely to function as gene silencers during Th1 differentiation. In contrast, all the remaining categories of DHS sites showed positive correlations with changes of gene expression during Th2 differentiation.

Discussion

Although many studies have analyzed chromatin remodeling or DHS sites at the genomic loci of Th1 and Th2 effector cytokine genes, there have been no studies that have analyzed genome wide chromatin remodeling, and how genome wide chromatin structures correlate with global gene expression in these cells and during their differentiation from naïve CD4 T cells. The present study fills these gaps of knowledge.

We used DNase-Seq to map genome wide distribution of DHS sites in naïve CD4 T cells, Th1 and Th2 cells. We first determined how the DHS sites are distributed genome wide. As the location of a DHS is important for determining whether the DHS site may regulate the expression of its associated gene, we divided the DHS sites into 6 location categories based on their locations relative to the gene bodies. We found that in all cell types, the genomic regions within 2kb surrounding the TSS of the associated genes have the highest densities of DHS sites. To lesser degrees, the regions from -5kb to -2kb and within 5kb downstream of TTS also have relatively high densities of DHS sites. In contrast, the region from +2kb to TTS and intergenic region have low densities

of DHS sites. These findings provided genome wide views of the chromatin structure landscapes of these cells.

For each DHS site, we assigned a DHS value to quantitatively measure its DNase I sensitivity or the “openness” of the chromatin. When all DHS sites were considered, there was a clear positive association between the DHS values and gene expression levels in all cell types. This result provided evidence that at the whole genome level the openness of the chromatin structure generally correlates positively with gene expression levels. When DHS sites in the different location categories were compared, highest positive association was found within 2kb of TSS, the same regions with the highest densities of DHS sites. For their proximity to the TSS, DHS sites in these regions may directly regulate transcription initiation, i.e., the promoter activities of the genes. In contrast, only marginal positive association was found between the numbers of DHS sites and gene expression when all DHS sites and their associated genes in a cell were considered. Nonetheless, comparatively the numbers of DHS sites from +2kb to TTS and the intergenic regions showed notably positive association.

Thus, it should be noted that there is a distinction of gene regulatory behaviors by DHS sites near the promoters and by those further downstream in the gene body and in the intergenic region. It appears that the numbers of DHS sites near or in the promoter regions are not dramatically different between genes of low and high expression levels. However, the DHS sites associated with genes of low expression levels are less open than those associated with genes of high expression levels. In contrast, chromatin in the +2kb to TTS and intergenic regions may be more tightly packed so that DHS sites are more likely to be associated with genes of high than low expression levels. In other words, whether DHS value or the number of DHS sites is the determinative factor for gene regulatory role depends on the genomic locations.

Our study also showed that during the differentiation from naïve CD4 T cells to Th1 or Th2 cells, chromatin remodeling is not just a local event restricted to the effector cytokine gene loci, but rather a genome wide event. The differentiated Th1 and Th2 cells have more open chromatin structures than naïve CD4 T cells. Changes of DHS values (Δ DHS) within 2kb of TSS showed the strongest positive correlation with changes of gene expression (Δ RPKM) in both Th1 and Th2 differentiation. In Th2 differentiation, positive associations between Δ DHS and Δ RPKM were also detected in all other location categories. However, in Th1 differentiation negative associations were found in the +2kb to TTS and intergenic regions, and to some degree in the -5kb to -2kb region. The negative correlations suggest that DHS sites in these regions are statistically more likely to function as gene silencers during Th1 differentiation. Consistent with this notion, it has been shown in an earlier study that a DHS site near the 3' untranslated region of the IL-4 gene acts as a silencer of IL-4 expression in Th1 cells (23). Thus, chromatin remodeling has different functional implications with regard to gene regulation in different cellular differentiation pathways. Findings in this study may serve as guides for experimental studies of *cis* regulation of gene expression during Th1 and Th2 differentiation.

Experimental Procedures

DHS library construction

Naïve CD4 T cells, Th1 and Th2 cells were prepared as previously described (24, 25). After cell culture, dead cells and debris were removed by centrifuge on LSM (MP Biomedicals, Santa Ana). Nucleus preparation and DNase I digestion were performed as described (26). DHS DNA libraries were constructed following the protocol developed in Dr. GE Crawford's lab (20, 27). For each cell type, at least 2 independent

replica libraries were constructed. Briefly, DNase I digested nuclei were embedded in low melt agarose gel. Genomic DNA were isolated and blunt-ended in gel, then ligated to a biotinylated MmeI adaptor. The adaptor-ligated high molecular weight DNA were gel purified and digested with MmeI. The MmeI digested DNA fragments containing 20bp of genomic DNA derived from the DHS regions were isolated using Streptavidin conjugated Dynal M-280 beads (Invitrogen). The isolated DNA fragments were ligated to adaptor II, which contains index sequence and sequencing primers. The ligation products were treated with 0.15N NaOH and washed extensively, then amplified with PCR primers. The 140 bp PCR products were the DHS libraries. Use of animals in this study was approved by the Institution Animal Care and Use Committee of Marshall University.

DNase Seq and data processing

The DHS libraries were gel purified and sequenced on an Illumina HiSeq 1000 genome analyzer at the Marshall University Genomic Core Facility. Adaptor sequences were trimmed from the sequence reads using Trimmomatics (28). The trimmed reads were aligned to the mouse reference genome mm10 using Bowtie (29). DHS peaks were called with F-Seq (21). Gene annotation of the DHS peaks was performed using the Perl script annotatePeaks.pl in HOMER (30). For intergenic DHS peaks, we compared the distances from a DHS peak to the TTS of its upstream gene and to the TSS of its downstream gene, then chose the gene that is closer to the DHS peak as the associated gene. For quantitative measurement, we used a formula similar to the previously described (22) to define the DHS value of a DHS site as $N_i / (\sum_{k=1}^m N_k) / m$, where N_i is the tag count in a 600bp interval centered on the i -th DHS peak and m is the total number of peaks in a cell type. Δ DHS values between 2 cell types were determined as previously described (22).

RNA-Seq data processing

RNA-Seq analyses of gene expression in naïve CD4 T cells, Th1 and Th2 cells were described previously (24). The RNA-Seq raw data files (2 replicas for each cell type) in FASTQ format were summarized in Table S3. All the data files passed quality control test by FastQC (31). The RNA-Seq reads were aligned to the mouse reference genome mm10 using TopHat (32). Gene expression quantification was performed using HTseq-count pipeline (33). We calculated the RPKM value for each gene using locally developed Perl code: $RPKM = (10^9 * C) / (N * L)$, C = number of reads mapped to a gene, N = total mapped reads in the experiment, L = exon length in base pairs for a gene. The average RPKM values for each gene between 2 replicates were used for further analysis.

Analyses of relationship between DHS sites and gene expression

Kendall tests were performed between the DHS values and the RPKM values for all or each of the location categories of the DHS sites using the Kendall R package (34) to calculate correlation coefficient or “tau” values. *P* values were obtained using the GeneNet R package (35) to determine whether the correlations were statistically significant. Relations between Δ DHS and Δ RPKM were analyzed similarly.

Figures and Tables

Figure 1. DHS libraries. Upper panel is the schematic diagram of the structure and size of the DNA fragments of the DHS libraries. Lower panel is the high resolution agarose gel picture of the final PCR products of the DHS libraries. pBR322/Msp I: molecular markers; nTcon: naïve conventional CD4 T cells; no DNase I: control DHS library derived from nuclei of total CD4 T cells not digested with DNase I.

Figure 2. Distribution of DHS sites (peaks) at distances to the associated genes. For panels B and C, the numbers of DHS peaks within continuous sliding windows of 100bp were plotted against the distance to the TSS or TTS. The maximum points of the X axes are arbitrarily set at the distances that contain 80% of total DHS sites. (A) Schematic diagram of the genomic regions covered in B and C. (B) Distribution of DHS peaks in region from -5kb to the TTS. (C) Distribution of DHS peaks in the regions downstream of TTS.

Figure 3. DHS peak densities in different genomic regions. Left panel, schematic diagram of 6 location categories of the genome as defined by their distances to gene bodies: -5kb to -2kb, -2kb to TSS, TSS to +2kb, +2kb to TTS, TTS to TTS + 5kb, and the intergenic region. Right panel, DHS peak densities (numbers of peaks/kb) in the 6 location categories of the genome. All (100%) DHS peaks in each cell type were included in the calculation.

Figure 4. Association between DHS values and gene expression levels (RPKM).

Genes of the transcriptome are divided into 4 groups based on their expression levels in a given cell – i.e., top 25%, top 25 - 50%, bottom 25-50% and bottom 25%. DHS values of the DHS sites associated with each of these groups of genes were calculated.

ANOVA tests were performed to calculate P values for each cell types, which are summarized in Table S1. (A) Relations between gene expression and the DHS values of all DHS sites. (B) Schematic diagram of the 6 location categories of the genome. (C) Relation between gene expression and DHS values of DHS sites in each of the 6 location categories. P values are provided in Table S1.

Figure 5. Association between the numbers of DHS sites and gene expression levels. Genes are divided into 4 expression groups as in Fig 3. (A) Relations between gene expression and the numbers of all associated DHS sites. (B) Schematic illustration of the 6 location categories of the genome. (C) Relations between gene expression and the numbers of the gene-associated DHS sites in each of the 6 location categories.

Figure 6. Dramatic changes of DHS during Th1 or Th2 differentiation. Change of DHS value (Δ DHS) between naïve CD4 T cells and Th1 or Th2 cells was calculated for each of the DHS sites identified in either pair of the cells. The numbers of DHS sites with increased, unchanged or decreased DHS values are shown.

Table 1. Correlation coefficients of DHS and RPKM values

Cell type	DHS peak location category	Value 1	Value 2	Correlation Coefficient	<i>P</i> value (GeneNet)
Naïve CD4	All categories	DHS	RPKM	0.13	0.00
	-5 to -2kb from TSS	DHS	RPKM	0.09	3.59E-15
	-2kb to TSS	DHS	RPKM	0.18	2.42E-73
	TSS to +2kb	DHS	RPKM	0.23	6.44E-173
	+2kb to TTS	DHS	RPKM	0.07	2.66E-75
	TTS to TTS + 5kb	DHS	RPKM	0.07	3.95E-06
	Intergenic	DHS	RPKM	0.10	9.38E-119
Th1	All categories	DHS	RPKM	0.12	0.00
	-5 to -2kb from TSS	DHS	RPKM	0.08	1.99E-10
	-2kb to TSS	DHS	RPKM	0.27	1.70E-178
	TSS to +2kb	DHS	RPKM	0.28	8.55E-283
	+2kb to TTS	DHS	RPKM	0.06	9.13E-58
	TTS to TTS + 5kb	DHS	RPKM	0.05	4.68E-03
	Intergenic	DHS	RPKM	0.07	6.83E-63
Th2	All categories	DHS	RPKM	0.13	0.00
	-5 to -2kb from TSS	DHS	RPKM	0.10	6.65E-18
	-2kb to TSS	DHS	RPKM	0.19	4.72E-90
	TSS to +2kb	DHS	RPKM	0.26	7.45E-260
	+2kb to TTS	DHS	RPKM	0.11	4.65E-188
	TTS to TTS + 5kb	DHS	RPKM	0.06	3.68E-04
	Intergenic	DHS	RPKM	0.09	7.08E-119

Table 2. Correlation coefficients of Δ DHS and Δ RPKM values

Cell type comparison	DHS peak location category	Value 1	Value 2	Correlation Coefficient	<i>P</i> value (GeneNet)
Naïve vs Th1	All categories	Δ DHS	Δ RPKM	4.53E-02	4.92E-152
	-5 to -2kb from TSS	Δ DHS	Δ RPKM	-1.06E-03	9.21E-01
	-2kb to TSS	Δ DHS	Δ RPKM	3.30E-02	3.37E-06
	TSS to +2kb	Δ DHS	Δ RPKM	3.40E-02	4.59E-09
	+2kb to TTS	Δ DHS	Δ RPKM	-2.31E-02	1.47E-16
	TTS to TTS + 5kb	Δ DHS	Δ RPKM	2.55E-02	2.96E-02
	Intergenic	Δ DHS	Δ RPKM	-6.43E-03	2.89E-02
Naïve vs Th2	All categories	Δ DHS	Δ RPKM	3.03E-02	9.74E-73
	-5 to -2kb from TSS	Δ DHS	Δ RPKM	2.83E-02	6.83E-03
	-2kb to TSS	Δ DHS	Δ RPKM	4.52E-02	1.89E-10
	TSS to +2kb	Δ DHS	Δ RPKM	4.74E-02	6.48E-17
	+2kb to TTS	Δ DHS	Δ RPKM	4.73E-02	5.20E-68
	TTS to TTS + 5kb	Δ DHS	Δ RPKM	3.88E-02	6.46E-04
	Intergenic	Δ DHS	Δ RPKM	2.82E-02	4.13E-21

Declaration of Conflict of Interest

The author declares no conflict of interest.

Author contribution

WZ designed, carried out the study, and wrote the manuscript.

References

1. Liew, F. Y. 2002. T(H)1 and T(H)2 cells: a historical perspective. *Nat Rev Immunol* 2: 55-60.
2. Stark, G. R., I. M. Kerr, B. R. Williams, R. H. Silverman, and R. D. Schreiber. 1998. How cells respond to interferons. *Annu Rev Biochem* 67: 227-264.
3. Finkelman, F. D., T. Shea-Donohue, S. C. Morris, L. Gildea, R. Strait, K. B. Madden, L. Schopf, and J. F. Urban, Jr. 2004. Interleukin-4- and interleukin-13-mediated host protection against intestinal nematode parasites. *Immunol Rev* 201: 139-155.
4. Annunziato, F., C. Romagnani, and S. Romagnani. 2015. The 3 major types of innate and adaptive cell-mediated effector immunity. *J Allergy Clin Immunol* 135: 626-635.
5. Jutel, M., and C. A. Akdis. 2011. T-cell subset regulation in atopy. *Current allergy and asthma reports* 11: 139-145.
6. Hsieh, C. S., S. E. Macatonia, C. S. Tripp, S. F. Wolf, A. O'Garra, K. M. Murphy, T. A. Wynn, D. Jankovic, S. Hieny, K. Zioncheck, P. Jardieu, A. W. Cheever, and A. Sher. 1993. Development of TH1 CD4⁺ T cells through IL-12 produced by Listeria-induced macrophages. *Science* 260: 547-549.
7. Seder, R. A., W.E. Paul, M.M. Davis, and B.F. de St. Groth. 1992. The Presence of Interleukin 4 During in vitro Priming Determines the Lymphokine-producing Potential of CD4⁺ Cells from T Cell Receptor Transgenic Mice. *J Exp Med* 176: 1091-1098.
8. Amsen, D., J. M. Blander, G. R. Lee, K. Tanigaki, T. Honjo, and R. A. Flavell. 2004. Instruction of distinct CD4 T helper cell fates by different notch ligands on antigen-presenting cells. *Cell* 117: 515-526.

9. Liew, F. Y., J. P. Girard, and H. R. Turnquist. 2016. Interleukin-33 in health and disease. *Nat Rev Immunol* 16: 676-689.
10. Zeng, W. P. 2013. "All Things Considered": Transcriptional Regulation of Th2 Differentiation from Precursor to Effector Activation. *Immunology* 140: 31 - 38.
11. Ziegler, S. F., F. Roan, B. D. Bell, T. A. Stoklasek, M. Kitajima, and H. Han. 2013. The biology of thymic stromal lymphopoietin (TSLP). *Adv Pharmacol* 66: 129-155.
12. Szabo, S. J., B. M. Sullivan, C. Stemmann, A. R. Satoskar, B. P. Sleckman, and L. H. Glimcher. 2002. Distinct effects of T-bet in TH1 lineage commitment and IFN-gamma production in CD4 and CD8 T cells. *Science* 295: 338-342.
13. Zheng, W.-p., and R. A. Flavell. 1997. The transcription factor GATA-3 is necessary and sufficient for Th2 cytokine gene expression in CD4 T cells. *Cell* 89: 587-596.
14. Wu, C. 1980. The 5' ends of Drosophila heat shock genes in chromatin are hypersensitive to DNase I. *Nature* 286: 854-860.
15. Djuretic, I. M., D. Levanon, V. Negreanu, Y. Groner, A. Rao, and K. M. Ansel. 2007. Transcription factors T-bet and Runx3 cooperate to activate Ifng and silence Il4 in T helper type 1 cells. *Nature immunology* 8: 145-153.
16. Gross, D. S., and W. T. Garrard. 1988. Nuclease hypersensitive sites in chromatin. *Annual review of biochemistry* 57: 159-197.
17. Agarwal, S., and A. Rao. 1998. Modulation of chromatin structure regulates cytokine gene expression during T cell differentiation. *Immunity* 9: 765-775.
18. Shnyreva, M., W. M. Weaver, M. Blanchette, S. L. Taylor, M. Tompa, D. R. Fitzpatrick, and C. B. Wilson. 2004. Evolutionarily conserved sequence elements that positively regulate IFN-gamma expression in T cells. *Proc Natnl Acad Sci USA* 101: 12622-12627.

19. Takemoto, N., N. Koyano-Nakagawa, T. Yokota, N. Arai, S. Miyatake, and K.-I. Arai. 1998. Th2-specific DNase-I-hypersensitive sites in the murine IL-13 and IL-4 intergenic region. *International Immunol* 10: 1981-1985.
20. Boyle, A. P., S. Davis, H. P. Shulha, P. Meltzer, E. H. Margulies, Z. Weng, T. S. Furey, and G. E. Crawford. 2008. High-resolution mapping and characterization of open chromatin across the genome. *Cell* 132: 311-322.
21. Boyle, A. P., J. Guinney, G. E. Crawford, and T. S. Furey. 2008. F-Seq: a feature density estimator for high-throughput sequence tags. *Bioinformatics* 24: 2537-2538.
22. He, H. H., C. A. Meyer, M. W. Chen, V. C. Jordan, M. Brown, and X. S. Liu. 2012. Differential DNase I hypersensitivity reveals factor-dependent chromatin dynamics. *Genome Res* 22: 1015-1025.
23. Ansel, K. M., R. J. Greenwald, S. Agarwal, C. H. Bassing, S. Monticelli, J. Interlandi, I. M. Djuretic, D. U. Lee, A. H. Sharpe, F. W. Alt, and A. Rao. 2004. Deletion of a conserved Il4 silencer impairs T helper type 1-mediated immunity. *Nature Immunology* 5: 1251-1259.
24. Hu, G., Q. Tang, S. Sharma, F. Yu, T. M. Escobar, S. A. Muljo, J. Zhu, and K. Zhao. 2013. Expression and regulation of intergenic long noncoding RNAs during T cell development and differentiation. *Nat Immunol* 14: 1190-1198.
25. Zeng, W. P., V. E. Sollars, and P. Belalcazar Adel. 2011. Domain requirements for the diverse immune regulatory functions of foxp3. *Molecular immunology* 48: 1932-1939.
26. Zeng, W. P., and M. M. McFarland. 2014. Rapid and unambiguous detection of DNase I hypersensitive site in rare population of cells. *PloS one* 9: e85740.

27. Song, L., and G. E. Crawford. 2010. DNase-seq: a high-resolution technique for mapping active gene regulatory elements across the genome from mammalian cells. *Cold Spring Harbor protocols* 2010: pdb prot5384.
28. Bolger, A. M., M. Lohse, and B. Usadel. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30: 2114-2120.
29. Langmead, B., C. Trapnell, M. Pop, and S. L. Salzberg. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10: R25.
30. Heinz, S., C. Benner, N. Spann, E. Bertolino, Y. C. Lin, P. Laslo, J. X. Cheng, C. Murre, H. Singh, and C. K. Glass. 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* 38: 576-589.
31. Simon, A. FastQC. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
32. Trapnell, C., L. Pachter, and S. L. Salzberg. 2009. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25: 1105-1111.
33. Anders, S., P. T. Pyl, and W. Huber. 2015. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31: 166-169.
34. McLeod, A. I. 2011. Kendall: Kendall rank correlation and Mann-Kendall trend test. <https://cran.r-project.org/web/packages/Kendall/index.html>.
35. Opgen-Rhein, R., and K. Strimmer. 2007. From correlation to causation networks: a simple approximate learning algorithm and its application to high-dimensional plant gene expression data. *BMC Syst Biol* 1: 37.

Supplemental Information

Figure S1. Representatives of DHS peaks of replica libraries. Each of the DHS replica libraries of naïve, Th1 and Th2 cells was sequenced and analyzed. DHS peaks were called using F-Seq. DHS peaks at the *CD4*, *IFN- γ* , and the *IL-4/IL13* loci of a representative library of naïve, Th1 and Th2 cells are shown, respectively.

Figure S2. Distribution of DHS sites at distances relative to gene body. The numbers of all DHS sites whose peak positions fall in the continuous sliding intervals of 100pb at the same genomic distances to their associated genes were counted. Histograms of the numbers of DHS sites and their distances to the associated genes are shown. (A)

Schematic diagram showing the genomic regions in which the DHS sites are located. (B) The distribution of DHS sites in the genomic regions from -5kb upstream of the TSS to the TTS of the associated genes. Distances are the numbers of bp to the TSS. (C) The distribution of DHS sites in the regions ranging from the TTS of the upstream gene to the -5kb of the downstream gene. Distances are the number of bp to the TTS.

Table S1. *P* values (ANOVA) in the analyses of the association between gene expression and DHS values

Genomic Location Category	Naïve CD4 T cell	Th1 Cell	Th2 Cell
All Sites	8.90E-42	1.13E-27	0.00E+00
-5kb to -2kb from TSS	2.83E-13	1.95E-11	1.09E-11
-2kb to TSS	1.96E-02	6.57E-05	5.96E-03
TSS to +2kb	3.42E-05	1.31E-10	1.80E-02
+2kb to TTS	2.16E-33	1.40E-09	2.06E-86
TTS to TTS + 5kb	5.57E-04	1.03E-02	1.93E-02
Intergenic	8.51E-02	8.54E-02	6.93E-05

Table S2. *P* values (ANOVA) in the analyses of the association between gene expression and the numbers of DHS sites

Genomic Location Category	Naïve CD4 T cell	Th1 Cell	Th2 Cell
All Sites	1.66E-148	2.04E-108	2.15E-44
-5kb to -2kb from TSS	1.65E-20	8.60E-03	2.67E-10
-2kb to TSS	7.12E-169	3.99E-297	6.23E-140
TSS to +2kb	1.32E-244	0.00E+00	5.65E-281
+2kb to TTS	5.35E-116	1.98E-09	1.93E-27
TTS to TTS + 5kb	5.29E-16	3.52E-01	2.93E-07
Intergenic	3.19E-33	7.14E-14	2.07E-23

Table S3. Summary of RNA-Seq data files

Accession #	Cell type	Total reads	Mapped reads*
GSM1169442	Naïve CD4	37,508,516	36,402,809
GSM1169444	Naïve CD4	33,635,005	31,948,385
GSM1169445	Th1	37,831,787	35,634,868
GSM1169446	Th1	37,030,979	35,782,352
GSM1169449	Th2	56,892,461	50,458,993
GSM1169450	Th2	61,688,729	48,074,057

*Adaptor sequences were trimmed from the reads. The trimmed reads were mapped to gene exons.